Hierarchical Markov Reward Model for Availability Analysis and Optimization of the Grid Computing System

Y.S. Dai, M. Xie

Department of Industrial & Systems Engineering,

National University of Singapore

ENGP0495@NUS.EDU.SG, MXIE@NUS.EDU.SG

Abstract Summary

The grid computing is a recently developed technique for complex systems with large-scale resource sharing, wide-area program communicating, and multi-institutional organization collaborating, e.g. Dai *et al.* (2002). The grid computing is highly controlled by the resource management system (RMS). Although the development tools for the grid have been widely investigated (Foster & Kesselman, 1998), the availability of the grid computing has not been specifically studied.

The grid availability is defined as the probability for the grid RMS to be available in providing services for the grid jobs. There are usually two cases that cause the unavailable grid: 1) All the servers are down; 2) the request queue is full so that no new job requests can be added in the queue. The grid availability is different from the availability of the conventional distributed systems (Lai *et al.*, 2002) that only considered the former type of unavailable case. Actually, the number of job requests in a global grid may be much great, especially when the OGSA (Open Grid Services Architecture) is applied, see e.g. Foster *et al.* (2002). The OGSA enables integration of services by various organizations and therefore the job requests for all these services are of a great number. Thus, the possibility of unavailable grid caused by full request queue is much more than that caused by the failed servers, especially during the peak time.

Hence, this paper originally presents a hierarchical Markov reward model for the grid availability analysis. Based on the model, an optimization problem for designing the grid system is studied. Then, the sensitivity analysis is conducted and a dynamic switching method is further suggested according to the sensitivity analysis.

References

Dai, Y.S., Xie, M. and Poh, K.L. (2002), Reliability Analysis of Grid Computing Systems, *IEEE Pacific Rim International Symposium on Dependable Computing (PRDC2002),* IEEE Computer Press, pp. 97-104.

Dai, Y.S., Xie, M., Poh, K.L., Liu, G.Q. (2003a), A study of service reliability and availability for distributed systems, *Reliability Engineering and System Safety*, vol. 79, no. 1, pp. 103-112.

Dai, Y.S., Xie, M., Poh, K.L., Yang, B. (2003b), Optimal testing-resource allocation with genetic algorithm for modular software systems, *Journal of System and Software,* vol. 66, no. 1, pp. 47-55.

Foster, I., Kesselman, C. (1998), *The Grid: Blueprint for a New Computing Infrastructure*, Morgan-Kaufmann.

Foster, I., Kesselman, C., Nick, J.M., Tuecke, S. (2002), Grid services for distributed system integration, *Computer*, vol. 35, no. 6, pp. 37-46.

Lai, C.D., Xie, M., Poh, K.L., Dai, Y.S., Yang, P. (2002), A model for availability analysis of distributed software/hardware systems, *Information and Software Technology,* vol. 44, no. 6, pp. 343-350.

Levitin, G., Dai, Y.S., Xie, M. and Poh, K.L (2003), Optimizing survivability of multi-state systems with multi-level protection by multi-processor genetic algorithm, *Reliability Engineering & System Safety*, vol. 82, pp. 93-104.